A publication of The Center on Online Learning
and Students with Disabilities

Learning Analytics

Center on Online Learning
and Students with Disabilities

Learning Analytics

inside.

What are the major topics that this White Paper includes?

- The Processes of Collection
- The Current Status of Use
- The Challenges
- Three Cautions

ANALYTICS

Online activities pervade 21st century daily life, from reading the news, to corresponding with friends and family, shopping, looking up driving directions, making reservations, filing tax returns, learning about subjects of individual interest, and so on. As individuals conduct these activities, they leave what one author referred to as "digital bread crumbs" (Brown, 2011). These crumbs are, in fact, data that can be analyzed to discern patterns and make predictions that can answer questions (e.g., who are the most likely buyers of a product?) or lead to real-time actions (e.g., ads displayed on a website). "Analytics" is the term currently used to describe the processes and technologies that continuously collect, analyze, and create meaningful visual representations of these digital data. Businesses, governments, academic institutions and others use these data representations to make informed decisions; draw reliable conclusions about current conditions and predict future events; or compare and improve the performance of individuals or groups (Norris, Baer, & Offerman, 2009).

Despite how common Analytics have become, many questions about them are unanswered and opportunities exist. For example, societal values and legal rights to the data generated by online activity and what ethically and legally may be done with these data are unsettled issues. Further, although sophisticated Analytics have been applied in some fields (e.g., commerce, national security), in other fields their potential has barely been tapped. This paper describes the current state and future challenges of Analytics in one such field, that is, K-12 education; and poses some unanswered questions particular to this context and population.

CHANGING APPLICATIONS

For nearly a decade, institutions of higher education have been developing and using Academic Analytics systems. The purpose of these systems is to support institutional management and decision making in relation to student enrollment, retention and graduation rate goals, and to demonstrate institutional accountability (Arnold, 2010; Campbell & Oblinger, 2007; Goldstein, 2005). Learning Analytics (LA) are the next wave in applying analytics in educational contexts, progressing from institutional goals and actions to individual student goals and actions (Hrabowski, Suess, & Fritz, 2011). LA has been defined as "the use of analytic techniques to help target instructional, curricular, and support resources to support the achievement of specific learning goals" (van Barneveld, Arnold, & Campbell, 2012, p. 8).

Computer-based learning environments have created opportunities for educators to shift from gathering data about learners after instruction to data collection, analysis and feedback or adaptations that optimize the learning environment during instruction, or in a broader sense, during any learning experiences (Ferguson & Buckingham Shum, 2012; Timms, Clements, Gobert et al., 2012). For example, LA tools may be integrated within online curriculum to collect data about each student's pace and accuracy with a set of learning tasks; pass these data through an algorithm based on research about student frustration levels; and determine the optimal pace and level of scaffolding the curriculum should present to each student for the next set of tasks. A series of papers presented at a recent conference describe the potential for LA to create such personal learning environments (PLE; Reinhart & Ullmann, 2011), which enhance learning by making students more aware about learning objects, providing them access to experts, giving them opportunities to connect with peers, and providing prompts from the system that alert students to news that emerges about a topic of interest.

"what ethically and legally may be done with these data are unsettled issues."

Not only can LA improve individual student learning experiences, teachers can improve their understanding about classes, cohorts or groups of learners. LA tools can prepare reports or visual representations of data that reveal patterns, trends and exceptions, triggering specific action such as scheduling a tutoring session with a struggling student (Brown, 2011). De Liddo and colleagues (2011) posit that LA tools enable educators to "interrogate the black box," that is, make sense of massive amounts of online data—some of which may not be directly observable—to get higher-order answers and make judgments about the significance and meaning of these data. Eventually, as LA predictive models based on these new sources of data improve, small quantities of evidence collected over a short time will help educators make more timely and even better predictions and decisions about student learning (Timms et al., 2012).

As the field develops, LA tools are moving from measuring, predicting, and providing feedback about knowledge acquisition and skill development (i.e., what students learn) to problem-solving processes, collaborative behaviors and processes, social networking, discourse, enquiry-based learning, curiosity, and disposition (i.e., how students learn).

LEARNING ANALYTICS PROCESS

Major stages of the LA process are: (a) select data; (b) capture data; (c) structure and aggregate data; (d) analyze data; and (e) represent data for use or sharing (Brown, 2011; Campbell & Oblinger, 2007; Elias, 2011; Hrabowski, Suess, Fritz, 2011; van Barneveld et al., 2012).

Select. So many data elements could be captured and analyzed that educators could virtually drown in the data (Snibbe, 2006). Therefore, LA is a goal-directed practice in which what needs to be known or predicted and by whom set the limits for which data to collect. Educators may want to know what parts of the online curriculum students use (exposure), how much time they spend in each part (engagement), what additional content they access (curiosity, help-seeking), the quality and level of their discourse about this content (critical thinking), how well are they progressing toward the specified learning goals (rate of growth), etc.

Capture. Given selected context and objectives, LA tools may collect or extract data from a single source or several sources in order to answer specific questions or prompt desired actions. Some of these data are by-products of online activity (e.g., logins, click counts, tagging, posting questions, linking to other sites); other data are intentionally created for educational assessment (e.g., answers to a quiz). In addition to the real-time online learning activities, LA data may include static data from administrative systems, such as student schedules, final grades, and demographic characteristics.

Structure and Aggregate. LA tools are designed to structure and aggregated the raw data they collect. For example, the University of Phoenix, an online learning program, has an LA tool that draws from 30 data sources for the purpose of predicting student persistence (Brown, 2011).

Some online learning data are naturally structured for use in LA (e.g., time elapsed for answering a question), and other data need to be transformed into analyzable structures. For example, De Liddo et al. (2011) are developing discourse-centered LA tools that can classify the rhetorical role a learner plays when posting in a wider online conversation (e.g., asks questions, contributes data, presents ideas). These tools also evaluate and classify the rhetorical moves in a posting (e.g., builds on an existing idea, solves a problem, refutes an argument, presents a challenge or counterpoint to a statement). Once structured into rhetorical roles and moves, these data, like the naturally structured data, can be included in analysis to predict, improve or evaluate student learning.

Analysis. By using computing power to analyze the enormous quantity of data from online learning activities, LA can expand accuracy and cognitive depth of inferences that educators make about where and how learning happens for their students (De Liddo et al., 2011). The multivariate statistical methods that might be employed range from fairly basic decision trees and regression analyses, to structured equation modeling (SEM), path modeling, principal components analysis, collaborative filtering algorithms, Bayesian networks, association rule mining, clustering, social network analysis, knowledge-based recommendation, machine learning, artificial intelligence, etc. (Corbitt, 2003).

LA analysis may be based on pre-determined algorithms that rely on previously understood relationships between learning behaviors or educational practices and learning outcomes. These algorithms would typically trigger an action by the system, the learner, or the teacher. Alternatively, an LA system may perform analyses that create new understanding about the relationships between online learning behaviors and learning outcomes and perhaps presents these to teachers or educational researchers.

Representations. LA often produces visual representations as well as tables, charts and other ways of displaying analyses to various audiences (e.g., students, teachers, administrators). Typically these representations will incite the intended audience to act on the new information. For a simple example, an online learning curriculum may use LA to generate a pie chart showing the student his proportion of correct and incorrect answers, perhaps suggesting the student seek help from another website or a teacher. For a more complex example, the system may visually present to a teacher a social network analysis depicting which of her students are central to and highly engaged in discourse on a social learning site, and which students are on the periphery or altogether disconnected from the conversation. Such a visual depiction may prompt teacher actions toward the disconnected students.

CURRENT STATUS OF LA

LA is an emerging field and does not yet have a common lexicon or conceptual framework (van Barneveld et al. 2012). However, a few international societies exist for the promotion of LA and similar concepts, and these organizations (e.g., Community for Advancing Discovery Research in Education; EDUCAUSE Learning Initiative; International Conference on Learning Analytics & Knowledge; Personal Learning Environment Conference) will likely create standards and practices in the near future.

A variety of tools are under development or in early stage use at this time. van Barneveld et al. (2012) reported on several Academic and/or Learning Analytics tools currently in use at Purdue University, Course Signals; University of Maryland, Check My Activity; University of Phoenix, Effectiveness Sources Portal (ESP); Capella University, Learning & Career Outcomes; University of Michigan, M-Reports Dashboard; University of California-San Diego, Sponsored Project Excellence Achieved through Redesign (SPEAR); ACT, Inc., Student

Readiness Inventory; and Sinclair Community College, Student Success Plan. These examples, while not exhaustive, demonstrate that institutions of higher education have been able to put into practice some of the concepts described here as well as in research papers and prototype designs. These early stage efforts imply the feasibility of using LA with K-12 populations and learning environments.

CHALLENGES FOR LA IN K-12 EDUCATION

Issues facing the field as LA migrates from higher education contexts to K-12 are many. Each stage of the LA process and its implementation in K-12 learning poses questions that are particular to the context and population. A few of these are presented below.

Select. Who will identify the particular objectives for elementary and secondary LA systems? Will online curriculum development be formed in response to high-stakes testing or with other educational priorities? Which pedagogy and research will guide the selected data and define the assumptions in analyses and algorithms? Will the tools possess construct validity and pedagogical validity? Can LA tools, or even should they, become the driving force for learning theory, pedagogy and student learning models?

Capture. Data capture also presents challenges that must be resolved for LA to become practical for widespread use. For all populations, but especially for education of minors, student privacy and permission to use their data must be considered. Who "owns" the exhaust trail of data, and who has the right to give permission to use it in improving individual student learning? In the past, researchers have been careful to obtain permission from parents and assent from students when using their performance data for research to create such knowledge. Even with proper permission, can or should schools trade student data for free LA tools that help commercial curriculum developers improve their systems and analyses?

Structure and Aggregate. As LA systems are implemented in many schools, districts, states and nations, will the needed data be generally measureable, reliably available and in formats and technologies that can be used by LA systems? Further, as developers define and represent different learning contents and processes, can student achievement be validly and reliably measured using readily available digital data? What are the optimal grain sizes from individual steps in learning to whole course or several courses to collect and aggregate?

Analysis. Which statistical methods will best classify and identify patterns in learning behavior, as well as predict or prompt new behaviors? Will LA have sufficient scale in data to answer discipline- or population-specific questions (Siemens & Long, 2011)?

Representations. How can LA avoid overwhelming students with recommendations and visualization that distract from rather than help with learning? Can LA tools have sufficiently universal design features to be helpful to students with disabilities? For educators, what inferences can be drawn about the meaning of visually represented patterns?

Implementation. Will K-12 schools and districts embrace a culture that uses LA for promoting meaningful learning? Teachers' roles will necessarily change when LA is integral to the educational process. How can schools introduce LA while continuing to value the role of teachers in the educational process? Will teachers perceive LA as a proxy for teacher evaluation processes and resist using the tools? What professional development will teachers need to make good interpretations of visually represented data and good decisions about their actions in response? Information Technology (IT) staff in schools will also experience role shifts from supporting to actually helping drive learning solutions.

FUTURE ISSUES FOR LA

Even as educational researchers, developers, and practitioners resolve the current questions posed above, innovations in technology and advancements in learning science will continue to present new issues for LA. For example, as new hardware input devices like cameras (facial and object recognition), graphic pads, e-pens, kinetic movement detectors, natural language processors widen the type of data possible to gather in the learning process, can LA tools incorporate data from different learning environments (e.g., games, museum exhibits, blended hands on/computer-based activities) beyond those confined to a particular online curriculum? As with LA and basic online learning, the field will need to know what to measure and how to interpret the data in order to make meaningful inferences about the learner/learning (Timms et al., 2012).

Another future issue for LA may be using LA with "discovery learning," wherein the learning objectives, participants and materials are not defined in advance, as in a social networking situation. Currently LA predictive analyses rely on assumed measures of success that are defined by teachers, curriculum or standardized assessments, and are therefore predictable.

An ongoing challenge for educators employing any new approach to learning or teaching, not just LA but certainly including LA is this: how to provide supportive experiences for students while continuing to experiment with and refine the new approach?

THREE CAUTIONS

Caution is always warranted when using output from any statistical model, but particularly the user-friendly visual representation in LA: association and correlation do not denote causation. Moreover, when LA predicts an outcome, teachers and students must not let those predictions become deterministic, that is, seal the outcome and stop trying to improve or change the outcome (Siemens & Long, 2011). Such a mindset would belie the very purpose of LA: to improve student learning. Lastly, Ferguson and Buckingham Shum (2011) caution that students must continue to develop their own meta-cognitive and learning-to-learn skills rather than relying solely on LA feedback. LA developers must somehow find the balance between encouraging independent ability for deep learning and scholarship with the temporary aid of LA feedback and supports.

REFERENCES

Arnold, K. (2010). Signals: Applying Academic Analytics. EDUCASE Quarterly, 33, 1-6.

Brown, M. (2011). Learning analytics: The coming third wave. Boulder, CO EDUCAUSE Learning Initiative. http://www.educause.edu/Resources/LearningAnalyticsTheComingThir/227287

Campbell, J., & Oblinger, D. (2007). Academic Analytics. Boulder, CO and Washington, DC: EDUCAUSE.

Corbitt, T. (November, 2003). Business intelligence and data mining. Management Services Magazine, 18.

De Liddo, A., Buckingham Shum, S., Quinto, I., Bachler, M. & Cannavacciuolo, L. (2011). Discourse-Centric Learning Analytics. Proc. 1st Int. Conf. Learning Analytics & Knowledge. Feb. 27-Mar 1, 2011, Banff. Eprint: http://oro.open.ac.uk/25829

Elias, T. (2011). Learning Analytics: Definitions, processes and potential. Creative Commons.

Ferguson, R. & Buckingham Shum, S. (2012). Social learning analytics: Five approaches. Proc. 2nd Int. Conf. Learning Analytics & Knowledge, (29 Apr-2 May, Vancouver, BC). ACM Press: New York. Eprint: http://oro.open.ac.uk/32910

Goldstein, P. (2005). Academic Analytics: The uses and management of information and technology in higher education. EDUCAUSE Center for Applied Research, Roadmap. http://www.educause.edu/ecar

Hrabowski, F., Suess, J, & Fritz, J. (2011). Assessment and analytics in institutional transformation. EDUCAUSE Review Archive, 46(5).

Norris, D., Baer, L. & Offerman, M. (September, 2009.) A national agenda for action analytics, National Symposium on Action Analytics, St. Paul, MN. http://lindabaer.efoliomn.com/uploads/settinganationalagendaforactionanalytics101509.pdf

Reinhart, W., & Ullmann, T. (Eds.). (July 11, 2011). Proceedings of the 1st workshop on awareness and reflection in personal learning environments. In conjunction with the PLE Conference 2011. Southampton, United Kingdom.

Snibbe, A.C. (2006). Drowning in data. Stanford Social Innovation Review, 39–45. http://www.ssireview.org/pdf/2006FA_feature_snibbe.pdf

Timms, M., Clements, D., Gobert, J., Jass Ketelhut, D., Lester, J., Reese, D., & Wiebe, E. (2012). New measurement paradigms. Community for Advancing Discovery Research in Education. http://cadrek12.org/sites/default/files/NMP%20Report%20041412_0.pdf.

Van Barneveld, A., Arnold, K. E., & Campbell, J. (2012). Analytics in higher education: Establishing a common language (White Paper). Boulder, CO: EDUCAUSE Learning Initiative. http://www.educause.edu/Resources/AnalyticsinHigherEducationEsta/245405